DOCUMENT RESUME

ED 171 757 TM 009 003

AUTHOR	Kolen, Michael J.; And Others
TITLE	Using Beta Distributions to Estimate Percentile Ranks and Accumulate Norms for Student Opinion of Teaching Items. Research Report No. 99.
PUB DATE	Mar 79
NOTE	61p.; Paper presented at the Annual Meeting of the American Educational Research Association (63rd, San Francisco, California, April 8-12, 1979)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	*Course Evaluation; Higher Education; *Rating Scales; *Statistical Analysis; *Student Evaluation of Teacher Performance
IDENTIFIERS	Likert Scales

ABSTRACT

Members of the beta family of distributions were used to estimate percentile ranks and to accumulate normative data collected in a university-wide system for gathering student opinions about teaching--including the areas of course content, objectives, instructor's behavior, teaching methods and materials, and outcomes of instruction. The fitted distributions were found to be acceptably accurate and to provide as accurate or more accurate percentile ranks than the linear interpolation method often used for this purpose. In addition, the estimation method investigated requires storage of only minimal data in order to accumulate (update) norms from year to year. The nature of the results suggests that the beta family may be a useful modeling distribution for norms or population distribution estimates on Likert-type items in other attitude and opinion research areas. (Author/MH)

USING BETA DISTRIBUTIONS TO ESTIMATE PERCENTILE RANKS AND ACCUMULATE NORMS FOR STUDENT OPINION OF TEACHING ITEMS

Michael J. Kolen
Douglas R. Whitney
William M. Smith

Research Report No. 99

Presented at the annual meeting of the
American Educational Research Association
San Francisco, California
April, 1979

Evaluation and Examination Service
The University of Iowa
Iowa City, Iowa

March, 1979

USING BETA DISTRIBUTIONS TO ESTIMATE

PERCENTILE RANKS AND ACCUMULATE NORMS FOR STUDENT

OPINION OF TEACHING ITEMS

Michael J. Kolen
University of Iowa

Douglas R. Whitney
American Council on Education

William M. Smith
University of Iowa

Abstract

The development of norms for many assessment and evaluation instruments sometimes involves the use of relatively few observations and the resulting empirical norms (e.g., percentiles) often show irregular "jumps" at one or both ends of the score scale. This paper reports on the use of mathematical models to improve on empirical norms in a situation where the empirical distributions can be expected to be highly skewed and for which few observations were initially available in the "light" tail.

Members of the beta family of distributions were used to estimate percentile ranks and to accumulate normative data collected in a university-wide system for gathering student opinions about teaching. The fitted distributions were found to be acceptably accurate and to provide as accurate or more accurate percentile ranks than the linear interpolation method often used for this purpose. In addition, the estimation method investigated requires storage of only minimal data in order to accumulate (update) norms from year to year. The nature of the results suggests that the beta family may be a useful modeling distribution for norms or population distribution estimates on Likert-type items in other attitude and opinion research areas.

USING BETA DISTRIBUTIONS TO ESTIMATE PERCENTILE RANKS AND ACCUMULATE NORMS FOR STUDENT OPINION OF TEACHING ITEMS

Questionnaires for collecting student opinions about teaching are used on most college campuses. Most systems for processing such ratings tabulate individual student ratings on each item and report class means (medians) to the instructor. Many systems also report percentile ranks so that the mean rating by a class on a particular item may be normatively compared to the mean ratings of an appropriate population of classes.

Two methods for arriving at the percentile ranks for class means for individual Likert-type student opinion of teaching items are compared in this paper. These methods are generally applicable in any situation in which norms for means on Likert-type items are desired. The first method investigated is the commonly used linear interpolation method in which the reported percentile ranks (or bands) are obtained via linear interpolation between data points in the empirical cumulative distribution of class means. In the second method, the parameters of a two parameter beta distribution are first estimated from the distribution of class means and then used to estimate the desired percentile ranks. The estimated beta distribution method was studied because, if found to be appropriate, it can be expected to provide a smooth, regular function for the percentile ranks, to require the storage of a fraction of the amount of information required with the linear interpolation method, and to provide a simple method for accumulating and updating the percentile norms as additional data are gathered. In addition, the estimates may be more accurate than those obtained in the usual way.

Linear Interpolation Method

To arrive at norms using either of the two methods, it is necessary to collect data on each of the items from a sample from the population or populations of interest. When using the linear interpolation method, the sample data (e.g., class means) are rank ordered and the percentile ranks for each of the data points are calculated from this empirical cumulative distribution. When the data collected are item means, which is the concern here, an unlimited number of different means (on a closed interval bounded by the scale values) could possibly occur. In order to arrive at the percentile ranks for means that were not included in the sample data, linear interpolation is generally used. [Note that other methods of interpolation such as the logistic method described by Marco (1977) could also be employed.] The linear interpolation procedure can be used to arrive at percentile ranks of item means collected in the future. Note, however, that all of the norming data (the entire empirical cumulative distribution) must be stored to update the norms using data collected in the future.

Since the number of classes included in the initial norming of student opinion of teaching items is usually "small" (typically fewer than 500), the function describing the relationship between the item class means and percentiles that is produced by the linear interpolation method is very irregular. When graphed, it often appears as a very jagged function--at least in some scale regions. In course ratings, the irregularities will tend to be more pronounced at intervals where the data are sparse and are especially notable at the lower tail of these (usually) negatively skewed distributions. It seems reasonable to assume that the "actual" (population) distribution function would be a smooth, regular function.

Estimated Beta Distribution Method

One way to arrive at a smooth and regular function is to estimate the

parameters of an appropriate family of theoretical, continuous distributions from which percentile ranks could be derived. If a particular [illegible] of distributions is appropriate, the resulting percentile ranks would [illegible] to be more accurate than those derived using the linear interpolation method.

Since the responses to the student opinion items investigated here were bounded [ranging from one (strongly disagree) to six (strongly agree)], a bounded family of distributions seemed most useful to adequately model the data. The use of an unbounded distribution, such as the [illegible], would assume that class means might range from minus to plus infini[illegible] an impossibility with Likert-type data, and necessarily introduce some [illegible] at both ends of the scale.

The beta family of distributions is bounded and can take on a wide variety of shapes. For this study, the lower bound of the beta[1] was fixed at a value of one (strongly disagree) and the upper bound at a value of six (strongly agree). The two necessary parameters estimated from the data allow the density function for each item to be symmetr[illegible] [illegible]hly skewed in either direction, or J-shaped, as well as other i[illegible] [illegible]ate shapes. Since the distribution of means on the student opini[illegible] [illegible] investigated here are typically bounded with skewed shape, the beta family would seem to be an ideal family of distributions to use as [illegible]del for the data.

The method chosen for estimating the two parameters of the beta distribution was the method of moments (Hogg and Craig, 1970). While a maximum likelihood procedure would probably provide for more accurate estimation of the beta parameters, the fact that it requires an iterative solution (Johnson & Kotz, 1970) and that it requires the storage of a large

[1]Actually, to achieve appropriate scale bounds, the probability density funtion was f(ax + b).

amount of information in order to accumulate norms over time, caused us to rule out maximum likelihood as a viable estimation procedure in the present situation. In circumstances permitting its use, however, some additional gain in accuracy might be expected.

The method of moments requires only three statistics (the number of class means, the mean of the class means, and the variance of the class means) to estimate the two beta parameters. The mean and variance of the class means are equated to the first and second moments of the beta distribution via algebraic formulas in order to estimate the two beta parameters (α and β).

Updating and Accumulating Norms

It is often desirable to accumulate and update norms over time. That is, as new data are collected they can be used in the estimation of percentile ranks. Assuming that the characteristics of the population of interest do not change substantially over time, the inclusion of additional data into the procedure for calculating percentile ranks would be expected to increase the precision of the estimated percentile ranks regardless of which method is used.

The accumulation of data and the updating of percentile ranks over time using the linear interpolation method requires that all of the previously collected data be stored and used in combination with the newly collected data. In order to accumulate data over time, using the estimated beta distribution method, it is necessary to store only the number, mean, and variance of the class means. As additional data are collected, the beta parameters and percentile norms can be re-estimated by considering only the number, mean, and variance of the previously collected data and by updating these values using the new data. Thus, if the estimated beta

distribution method is found to be appropriate, it would require the storage of only a relatively small amount of information for each item and provide a simple method for accumulating percentile norms as desired.

METHODS FOR ESTIMATING AND ACCUMULATING PERCENTILE RANKS

Linear Interpolation Method

Estimation. The percentile rank of an item mean [PR($\overline{X}$)] contained in the initial norming data set was calculated as,

$$PR(\overline{X}) = \text{Pct Below } \overline{X} + \frac{1}{2}(\text{Pct at } \overline{X}). \qquad (1)$$

Linear interpolation was used to estimate the percentile ranks for item means between those item means contained in the initial norming data set.

Accumulating data. When new data are encountered, the percentile ranks from the initial norming data could be used to estimate the percentile ranks for the new data. If a sizable number of new means were collected, it might be desirable to include the new data to produce an updated set of estimated percentile ranks. That is, as new data are accumulated, they could be combined with the previously collected data to produce updated percentile ranks. Note that this requires the storage of all new and previously collected data.

Estimated Beta Distribution Method

Estimation. The formula of the beta density function of y, where y ranges from zero to one is,

$$f(y) = \frac{1}{B(\alpha,\beta)}\, y^{\alpha-1}(1-y)^{\beta-1}(0 \le y \le 1). \qquad (2)$$

In this formula α and β are the parameters which define the shape of the beta density function and were estimated via the method of moments in the present study. The $B(\alpha,\beta)$ expression is the beta function (Novick & Jackson, 1974) and is a scaling constant.

The α and β parameters as estimated by the method of moments for a variable ranging from zero to one are

$$\hat{\alpha} = M^2(1-M)/V-M \qquad (3)$$

amd

$$\hat{\beta} = M(1-M)/V-\hat{\alpha}-1 \qquad (4)$$

In these equations, M is the sample mean and V the sample variance of a variable ranging from zero to one.

Although a more general form of the beta density function which includes upper and lower bound parameters (or constants) exists, the form of the beta density function in (2) was used for simplicity. To use (2) in the present study, it was necessary to subtract one from each class mean and divide the resulting value by five in order to relocate the class means on a [0,1] interval. The mean and variance of the relocated values were used to estimate α and β by (3) and (4). Percentile ranks of the relocated means were estimated by integrating (2) over the appropriate limits and multiplying the resulting value by 100. The integration was completed with the locally developed program listed in the Appendix although the IMSL (1978) packaged routine, MDBETA, could have also been used. (The locally developed program was used because it was found to be more efficient.) The zero to one scale was then transformed back to a one to six scale while keeping the associated percentiles intact.

Accumulating data. The only summary statistics from the initial norming which are stored in order to accumulate data and produce updated percentile ranks are the number of class means, the mean of the class means, and the variance of the class means from the initial norming sample. The number of class means, the mean of the class means, and the variance of the class means for a new sample can be pooled with the statistics from the initial norming to produce an updated number of class means, mean of class

means, and variance of class means. Updated α and β parameter estimates can then be calculated using (3) and (4) with the updated statistics and the updated percentile ranks can be calculated using the updated estimates of α and β.

The updating procudures for the two methods were not evaluated in the present paper except for their relative ease of implementation. The procedures have been included for the purpose of indicating how they might be accomplished.

EVALUATION METHOD

Data Source

The data for the initial beta fitting and comparisons (the initial sampling) were gathered by administering each of 80 student opinion of teaching items in a representative sample of 189 classes at The University of Iowa in Spring, 1974. The 80 items were sampled from the general areas of course content, objectives, and structure; instructor's behavior; instructional methods and materials; and outcomes of instruction. Each of the items was scored on a Likert scale ranging from one (strongly disagree) to six (strongly agree).

Fourteen of the initial norming items were administered in a different, but comparable, sample of 210 classes at The University of Iowa in Fall, 1975. These data were used as a cross-validation check of both of the percentile estimation methods.

Details regarding the methods used to secure participation of instructors, documentation of the representativeness of the sample, and items used are provided in two local publications for interested readers [Evaluation and Examination Service Summary Report No. 44 (1976) and Evaluation and

Examination Service Memo No. 26 (1978)]. The numbers used to identify the items in the present paper correspond to the item identification numbers in the latter report.

Procedure

Percentile ranks for each of the 80 items included in the initial norming study were estimated using each of the two methods. The fit of the beta to the initial sample data was evaluated using a Kolmogorov (Conover, 1971)-like statistic. Let beta $(\overline{X})$ be equal to the value of the beta (cumulative) distribution function at $\overline{X}$ and $\hat{P}(\overline{X})$ the percentile rank at $\overline{X}$ defined in (1) and divided by 100. The statistic T was defined as,

$$T = \underset{X}{\text{supdif}} \mid \text{beta}(\overline{X}) - \hat{P}(\overline{X}) \mid .$$

That is, for those means contained in the initial norming sample, T is the greatest absolute difference between the percentile rank of a mean computed from (1) and the percentile rank of a mean estimated using the estimated beta distribution method divided by 100. Note that T is not a true Kolmogorov statistic and that the critical values used in this paper are probably too large to fully control the significance levels at the intended values[1].

Two additional Kolmogorov-like statistics were calculated for each of the 14 items included in the cross-validation sample. The statistics were designed to evaluate which of the two methods produced percentile ranks that compared most closely to percentile ranks which were calculated by applying (1) to the cross-validation norming data only. That is, the

---

[1]The statistic T is not a Komolgorov statistic because percentile ranks instead of observed cumulative distribution values were used, because there were a large number of "tied" means, and because the beta parameters were estimated from the data. Noether (1967) has indicated that a large number of ties leads to a conservative test. The use of the data to estimate the beta parameters would also be expected to lead to a conservative test.

statistics calculated from the cross-validation data were used as a criterion to evaluate the relative accuracy of the two methods.

For the cross-validation data, let $\overline{X}_c$ be a mean observed in the cross-validation sample and $\hat{P}_c(\overline{X}_c)$ be the percentile rank of $\overline{X}_c$ divided by 100 that was calculated from the observed cross-validation data using (1). Let $beta_I(\overline{X}_c)$ be the value of the beta distribution function which was estimated from the initial norming data and evaluated at $\overline{X}_c$. Let $\hat{P}_I(\overline{X}_c)$ be the percentile rank of $\overline{X}_c$ estimated using the linear interpolation method with the initial norming data. The test statistics are then,

$$T_1 = \underset{\overline{X}_c}{\text{supdif}} \left| \hat{P}_I(\overline{X}_c) - \hat{P}_c(\overline{X}_c) \right|,$$

and

$$T_2 = \underset{\overline{X}_c}{\text{supdif}} \left| beta_I(\overline{X}_c) - \hat{P}_c(\overline{X}_c) \right|.$$

For the means observed in the cross-validation sample, T_1 and T_2 are the greatest absolute difference between the percentile rank of all of those means estimated from the cross-validation sample only and the percentile rank of that same mean estimated using the linear interpolation and beta distribution on the initial sample data after the percentile ranks are divided by 100.

Smaller values of T_1 and T_2 indicate a closer correspondence between the percentile ranks estimated from the initial sample by the respective method and the percentile ranks observed in the cross-validation. The linear interpolation method will be judged to provide for a more accurate representation of the cross-validation data for those items where T_1 is smaller than T_2. The estimated beta distribution method will be judged to provide for a more accurate representation of the cross-validation data for those items where T_2 is smaller than T_1. The values of the summary statistics T_1 and T_2 were compared to one another and not to Komolgorov

critical values.

Goodness of fit comparisons considered thus far apply only to those situations in which the initial norming sample size is near 189. To allow greater generalizability of results, random samples of the class means from the initial sample were drawn for each of the 14 items that were also included in the cross-validation norming. For each of the 14 items, ten random samples were drawn with each of the probabilities of selection of .10, .25, .50, .75, and.90 for each class mean observed in the initial norming. The linear interpolation and estimated beta distribution methods were used to estimate percentile ranks and the T_1 and T_2 statistics were calculated using the cross-validation data.

RESULTS

Initial Norming Data

The number of means, the mean and variance of means, the $\hat{\alpha}$ and $\hat{\beta}$ values for the estimated beta distribution method, and the value of the test statistic T are presented in Table 1. The item numbers in Table 1 correspond to the item identification numbers listed in the Evaluation and Examination Service Memo No. 26 (1978).

Those items for which the T statistic surpassed the $\alpha=.05$ critical value for a two-tailed Kolmogorov statistic are noted. Had the Kolmogorov test been strictly appropriate, four of the 80 T statistics (80 X .05 = 4) would be expected to surpass the critical value if each of the beta distributions fit the corresponding population data. Since the Kolmogorov test is probably conservative in the present situation, fewer than four of the T statistics would be expected to surpass the $\alpha=.05$ Kolmogorov critical value if each of the beta distributions fit the corresponding data. As noted in Table 1, eight of the T statistics surpassed the $\alpha=.05$

Kolmogorov critical value. This indicates that the beta family did not fit the population distribution for some items. However, the fit appears to be reasonable with the median value of T being .0715. That is, the greatest absolute difference between the observed percentile ranks and estimated beta distribution method percentile ranks, after the percentile ranks were divided by 100, was no more than .0715 for half the items and no more than .1346 for any item in the set.

Insert Table 1 About Here

Cross-Validation Data-Complete Sample

The number of means, the mean and variance of the means, and the T_1 and T_2 statistics for the cross-validation norming are presented in Table 2. Unfortunately, the interpretation of these results is ambiguous. The T_1 statistic was smaller than the T_2 statistic for 10 of the 14 items ($p<.18$ for a two-ended sign test) suggesting that the linear interpolation method may be more accurate. However, the median T_1 statistic (.0977) was larger than the median T_2 statistic (.0944) suggesting that the estimated beta distribution method was more accurate. Thus, we conclude that when the initial norming sample size was in the neighborhood of 189, both methods appeared to be nearly equally accurate for estimating percentile ranks for the student opinion of teaching items studied here.

Insert Table 2 About Here

Graphical Comparisons. Graphs of the fitted beta, the initial norming percentile ranks and interpolated percentile ranks, and the cross-validation data for the 14 items included in the cross-validation norming are provided

in Figures 1-14[2]. Overall, these figures illustrate the smoothness and regularity of the linear interpolated graphs. Visual inspection reveals that the fit of the beta to the initial norming data was probably as good as could be expected from a class of bounded curves for which only two parameters are estimated from the data. Thus, the beta appears to be an acceptable modeling distribution for the student opinion of teaching items which were studied.

Insert Figures 1-14 About Here

Evaluation of the differences between the two methods may be facilitated by referring to the graphs for which the T_1 and T_2 statistics differed by the greatest margin. The item (number 108) for which the difference between T_1 and T_2 was the largest and favored the linear interpolation method is shown in Figure 1. The percentile ranks produced by the two methods for item 108 differed by the greatest amount in the neighborhood of a scale value of 5. The cross-validation norming data was closer to the linear interpolated percentile ranks than to the estimated beta distribution percentile ranks in this interval. It appears to be the case that with this item, the estimated beta distribution method smoothed out an apparent irregularity which was actually a real bend in the distribution function. This illustrates, of course, the danger in assuming the degree of irregularity implied in the beta model. The beta cannot provide

[2]The linear interpolation method results are represented by a series of 'o' symbols and interpolation between points indicated by a solid line. The estimated beta distribution method results are represented by a smooth solid line and the cross-validation results are represented by a series of '+' symbols.

for an adequate fit to this item because the actual distribution function apparently has more than one point of inflection.

The data for the item (number 241) for which the difference between T_1 and T_2 was the largest and favored the estimated beta distribution method are graphed in Figure 6. For this item, the percentile ranks produced by the two methods differed most in the neighborhood of a scale value of 5. The "jump" in the graph of the linear interpolation method for the initial norming data occuring near 5 was smoothed out by the estimated beta distribution method. For this item, a large irregularity present with the linear interpolation method was presumably due to sampling error because the irregularity was not present in the cross-validation data and was the reason that the estimated beta distribution method was found to be substantially more accurate.

The data for items 108 and 241, and to a lesser extent the data for the other items, illustrated in Figures 1-14 tend to support the conjecture that the estimated beta distribution method smooths out irregularities and in cases where those irregularities are due to sampling error, the estimated beta method appears superior. In other cases, the beta seems to smooth out apparent irregularities which were also present in the cross-validation data, and thus, may not be due to sampling error. If this conjecture holds, a procedure for identifying those items for which the estimated beta distribution may not provide adequate estimates would be desirable. Items identified by this procedure could be closely monitored and, if necessary, the linear interpolation method used for estimating and accumulating percentile norms for these items. A method based on cross-validation data could be used for the assessment although no specific method is proposed.

The analysis of the data with the complete set of 189 means from the initial sample suggests that each method was better in some instances but that, overall, the methods produced nearly equally accurate results. These data do not indicate which method will be expected to be most accurate when other sample sizes or other kinds of items are used. The next set of results address this question of the comparative accuracy of the methods for smaller sample sizes.

Cross Validation Data--Small Samples

The median T_1 and T_2 summary statistics for each of the 14 items at each sample size as well as the median T_1 and T_2 summary statistics over all of the items are presented in Table 3. As expected, there was a tendency for the accuracy of both of the methods to decrease as the sample size decreased. This tendency was especially pronounced at the smallest sample sizes used in the study ($N \simeq 19$ and 49 classes).

Consistent with the results from the full sample sizes presented in Table 2, the median value of T_1 over all items was greater than the median value of T_2 over all items at each sample size. This result provides one indication that the estimated beta distribution method produced more accurate percentile norms than did the linear interpolation method.

Insert Table 3 About Here

The proportion of the 10 samples at each item and sample size combination as well as the proportion over all items at each sample size for which the individual (not median) value of T_1 was greater than T_2 is also shown in Table 3. Note that these values indicate the proportion of samples for which the estimated beta distribution method provided a closer fit to the cross-validation data than did the linear interpolation method.

For smaller sample sizes (say 50 or fewer class means), the estimated beta distribution results corresponded more closely to the cross-validation data than did the results from the linear interpolation method, but this does not necessarily hold for larger sample sizes. This conclusion holds whether the median values of T_1 and T_2 or the proportion of times that the T_1 statistic was greater than the T_2 statistic was used to evaluate fit.

In order to be able to make more precise statements, an analysis of variance was completed on the individual T_1 and T_2 values. The main effects were item (14 levels) which was considered as a between random effect, sample size (5 levels) which was considered as a between fixed effect, method (2 levels) which was considered as a within fixed effect, and sample (10 levels) which was considered as the random "subjects" effect. The ANOVA summary table, including estimated variance components, is shown in Table 4. The notation used to indicate main effects and interactions is that of Myers (1972). All effects and interactions tested surpassed the .01 critical value except for the methods main effect. Since item and sample were considered to be random effects, their main effects and interactions will not be discussed further.

The size, method, size x method, and overall means are shown in Table 5. Of special interest is the size x method interaction. Paired comparisons of method means at each level of sample size indicate that the methods differed only at the two smallest sample sizes. The difference was in favor of the estimated beta distribution method. Thus, with sample sizes of around 50 or fewer classes, the estimated beta distribution method was found to be superior. For the larger sample sizes studies, the two methods produced results which were about equally accurate.

DISCUSSION

The estimated beta distribution and linear interpolation methods for estimating percentile norms were found to provide nearly equally close approximations to the observed percentile ranks of the cross-validation data for the large initial sample sizes. For initial samples of 50 or fewer class means, the estimated beta distribution method was found to produce more accurate results than the linear interpolation method.[3]

For smaller sample sizes, the greatest accuracy was achieved by the estimated beta distribution method. Even for the larger sample sizes which were studied, the estimated beta distribution produced reasonably accurate results which were judged to be as accurate as the results produced by the linear interpolation method. Thus, the estimated beta distribution method would be expected to produce reasonably accurate results for the kinds of items from which the student opinion of teaching items included in this study were selected.

In addition to providing reasonably accurate results, the estimated beta distribution method may prove less costly and easier to use than the linear interpolation method in many situations. The estimated beta distribution method requires the storage of only a small fraction of the amount of information required with the linear interpolation method when the percentile norms are to be used in the future. For example, the University of Iowa cafeteria system for collecting student opinion of teaching data contains 200 items. If there were, on the average, 100 class means for

[3]It would be expected that, unless the beta model fit the population data perfectly (which is unlikely), at some sample size greater than 189 the linear interpolation method would produce more accurate results because it allows for a much wider (nearly infinite) range of forms for the function relating percentiles to percentile ranks.

for each item, then 20,000 (# items X mean # class means) of data would need to be stored and analyzed each semester if the linear interpolation method were used. As more class means were used to update the percentile norms, even more than 20,000 pieces of data would need to be stored and analyzed, even if the number of items remained constant. In contrast, only 600 (# items X 3) pieces of data would need to be stored with the estimated beta distribution method. This figure would not increase unless more items were added to the item pool. Note that rounding the item means to, say, the nearest tenth or using stanines could decrease the storage requirements for the linear interpolation method. However, this would still not reduce the storage requirements to three summary statistics per item, the number required with the estimated beta distribution method.

CONCLUSION

The beta family of distributions was found to provide reasonably accurate estimations of percentile ranks for class means arising from the use of Likert-type student opinion of teaching items. The estimated beta distribution method, in addition to providing as accurate or more accurate norms, produces a smoother function, requires the storage of less information, and provides for a method for updating and accumulating norms which requires substantially less storage of data than the linear interpolation method. Based on these considerations, we are presently using the estimated beta distribution method for creating, updating, and accumulating norms for the student opinion of teaching items which are administered at the University of Iowa. The estimated beta distribution method is especially recommended for accuracy purposes with sample sizes less than 50 class means as well as for simplicity and cost reasons when percentile norms are to be

updated and accumulated.

The estimated beta distribution offers the potential for improving both the accuracy and the economy of the student rating summary results reported to faculty members. Since these systems are ofte y used and results are frequently considered in critical promotion, nd tenure decisions, improvements in either respect over the methods c itly used could be of substantial importance in the retention of faculty seen (by students) to be good or poor teachers and in the modification of teaching methods based on student opinions. More generally, the close fit of the beta to the data indicates that the beta should be investigated as a possible model for aggregate statistics (e.g., means, medians) arising from the use of Likert-type items in other applications.

References

Conover, W. J. Practical nonparametric statistics. New York: Wiley, 1971.

Evaluation and Examination Service Memo No. 26. Students perceptions of teaching (SPOT): III. The 1978-79 item pool. University of Iowa, 1978.

Evaluation and Examination Service Summary Report No. 44. Students perceptions of teaching (SPOT): IV. Baseline data for a representative sample of classes. University of Iowa, 1976.

Hogg, R.V. & Craig, A.T. Introduction to mathematical statistics. 3rd. Ed. New York: MacMillan, 1970.

IMSL Library 1, (Fortran IV) IBMS/370-360. 7th Ed. Houston: International Mathematical and Statistical Libraries, Inc., 1978.

Johnson, N.L. & Kotz, S. Continuous univariate distributions-2. Boston: Houghton Mifflin, 1970.

Marco, G. L. Use of the logistic model as an alternative to linear interpolation for computing percentile ranks. Journal of Educational Measurement, 1977, 14, 271-276.

Noether, G.E. Elements of nonparametric statistics. New York: Wiley, 1967.

Novick, M.R. & Jackson, P.H. Statistical methods for educational and psychological research. New York: McGraw-Hill, 1974.

TABLE 1

Item Statistics For Class Means Initial Norming Study (N = 189 classes)

Item	Sample Statistics Mean	Variance	Estimated Beta Parameters Alpha	Beta	T#
101	4.64	0.70	4.41	1.65	0.1034*
102	4.60	0.77	4.01	1.56	0.0797
103	4.88	0.61	4.77	1.37	0.0588
104	5.06	0.55	4.81	1.12	0.1057*
108	4.66	0.74	4.09	1.50	0.0628
111	5.06	0.33	8.45	1.96	0.0504
112	4.54	0.61	5.24	2.16	0.1014*
114	4.63	0.57	5.56	2.10	0.0738
115	4.72	0.60	5.19	1.79	0.1129*
116	4.91	0.46	6.42	1.80	0.0785
117	4.83	0.76	3.73	1.13	0.0900
122	4.62	0.58	5.51	2.11	0.0699
123	4.72	0.52	6.03	2.08	0.1043*
124	4.85	0.70	4.31	1.28	0.0462
127	5.16	0.56	4.35	0.88	0.0738
128	4.72	0.74	4.07	1.40	0.0713
129	4.91	0.47	6.33	1.77	0.0967
130	5.27	0.33	7.35	1.26	0.1027*
201	4.83	0.62	4.79	1.46	0.0782
203	4.26	0.84	3.74	1.99	0.0469
206	4.62	0.65	4.88	1.86	0.0572
207	5.00	0.37	7.81	1.95	0.0535
208	5.31	0.44	5.01	0.81	0.0726
209	5.43	0.36	5.29	0.67	0.0696
210	4.82	0.66	4.41	1.36	0.0595
211	4.56	0.78	3.94	1.60	0.0640
212	4.81	0.54	5.64	1.76	0.0668
214	5.17	0.42	6.12	1.22	0.0715
215	4.74	0.60	5.15	1.74	0.0780
216	4.85	0.50	5.99	1.80	0.0881
217	5.21	0.59	3.91	0.73	0.0866
218	4.64	0.73	4.18	1.56	0.0543
219	4.64	0.72	4.27	1.59	0.0595
221	5.10	0.52	5.05	1.12	0.0682
226	5.04	0.51	5.27	1.25	0.0605
228	5.10	0.54	4.78	1.05	0.0805
233	4.58	0.61	5.25	2.08	0.0782
234	4.52	0.79	3.92	1.65	0.0717
239	5.02	0.49	5.65	1.37	0.0759
241	4.84	0.47	6.57	1.99	0.1346*
242	4.84	0.73	3.92	1.19	0.0455
245	4.70	0.87	3.37	1.18	0.0417
247	4.90	0.68	4.13	1.17	0.0809
301	4.90	0.58	4.94	1.38	0.0775
303	4.24	0.82	3.85	2.09	0.0897
304	4.68	1.07	2.59	0.92	0.0783
305	4.33	1.00	3.05	1.53	0.0540

TABLE 1 (cont.)

Item	Sample Statistics Mean	Variance	Estimated Beta Parameters Alpha	Beta	$T^{\#}$
306	4.59	0.95	3.09	1.21	0.0847
309	4.91	0.72	3.82	1.07	0.0620
312	4.31	0.86	3.66	1.88	0.0551
315	4.81	0.68	4.33	1.35	0.0589
317	5.17	0.31	8.51	1.69	0.0988
318	5.12	0.33	8.20	1.76	0.0696
321	4.86	0.66	4.35	1.28	0.0941
322	4.78	0.45	6.96	2.24	0.0918
326	4.64	0.75	4.04	1.50	0.0465
329	4.37	1.02	2.96	1.43	0.0611
332	4.77	0.73	4.07	1.33	0.1190*
333	4.55	0.50	6.53	2.66	0.0664
402	4.94	0.68	4.10	1.10	0.0781
403	4.28	0.97	3.18	1.67	0.0477
404	4.04	0.99	3.05	1.97	0.0494
405	4.44	0.75	4.27	1.93	0.0838
407	5.26	0.30	8.10	1.42	0.0857
408	4.99	0.59	4.68	1.19	0.0754
409	4.46	0.55	5.99	2.66	0.0734
410	4.32	0.67	4.85	2.46	0.0704
411	4.96	0.47	6.19	1.63	0.0679
413	4.73	0.46	7.02	2.40	0.0738
414	4.32	0.76	4.20	2.13	0.0782
415	4.75	0.46	6.95	2.31	0.0546
416	4.87	0.43	7.10	2.06	0.0588
417	4.38	0.57	5.78	2.78	0.0489
418	4.46	0.80	3.90	1.73	0.0529
420	4.60	0.66	4.74	1.84	0.0582
423	4.51	0.66	4.82	2.04	0.0798
424	4.72	0.66	4.61	1.58	0.0568
425	4.69	0.57	5.54	1.96	0.0798
427	4.60	0.68	4.60	1.79	0.0660
428	4.23	0.94	3.27	1.79	0.0432

*Indicates lack of fit of the estimated beta distribution method estimate to the results from the linear interpolation method at approximated α=.05.

#T is the greatest absolute difference between the percentile rank of a mean from the initial sample data using the linear interpolation method vs. the estimated beta distribution method.

TABLE 2

Summary Statistics and Test Statistics
for Cross-Validation
(N = 210 classes)

Item	Cross-Validation Statistics		Test Statistics	
	Mean	Variance	T_1	T_2
108	4.62	0.57	.0798	.1104
203	4.19	0.70	.0851	.1017
207	4.67	0.57	.1995	.2194
209	5.46	0.28	.0945	.0629
221	4.91	0.51	.1552	.1767
241	4.86	0.51	.1139	.0389
245	4.78	0.62	.0778	.0852
306	4.63	0.79	.0638	.0678
402	4.83	0.54	.1246	.1430
405	4.40	0.62	.1080	.0667
407	5.06	0.26	.2139	.2341
415	4.80	0.35	.0969	.0671
416	4.87	0.39	.0791	.0883
417	4.26	0.57	.0985	.1004
Median			.0977	.0944
P[1]			.29	

[1]Proportion of the 14 items for which $T_1 > T_2$, that is for which the estimated beta distribution method was superior.

TABLE 3

Median Values of T_1 and T_2 at each of Five Sample Sizes and Proportion of Times $T_1 > T_2$

Item	.90(N≈170)			.75(N≈142)			.50(N≈94)			.25(N≈47)			.10(N≈19)		
	T_1	T_2	P^1	T_1	T_2	P^1	T_1	T_2	P^1	T_1	T_2	P^1	T_1	T_2	P^1
108	.0820	.1128	.00	.0831	.1017	.00	.0998	.1144	.10	.1373	.1422	.30	.2236	.1187	.50
203	.0915	.1122	.00	.0870	.1063	.00	.1048	.1110	.30	.1394	.1150	.30	.2154	.1158	.60
207	.2068	.2217	.00	.2023	.2129	.00	.1937	.2184	.20	.2180	.2254	.40	.2446	.2118	.90
209	.0894	.0610	1.00	.1008	.0640	1.00	.0928	.0615	1.00	.1023	.0764	.70	.1573	.1018	.70
221	.1626	.1747	.00	.1606	.1749	.00	.1618	.1802	.10	.1583	.1676	.30	.2124	.2187	.60
241	.1155	.0382	1.00	.1027	.0428	1.00	.0920	.0542	.90	.1276	.0646	1.00	.1233	.1084	.90
245	.0813	.0870	.20	.0882	.0878	.20	.0848	.0866	.50	.1093	.1189	.20	.1754	.1391	.90
306	.0657	.0720	.10	.0718	.0765	.30	.0843	.0909	.30	.1236	.0715	.90	.1371	.1277	.50
402	.1235	.1380	.00	.1050	.1262	.00	.1211	.1450	.20	.1139	.1335	.30	.1908	.1825	.50
405	.1046	.0742	1.00	.1084	.0759	.90	.0881	.0641	1.00	.1130	.0991	.50	.1852	.1321	.70
407	.2151	.2361	.00	.2284	.2463	.10	.1974	.2246	.20	.2820	.2486	.40	.2168	.2200	.40
415	.1052	.0676	1.00	.1151	.0757	.90	.1029	.0808	.90	.1297	.1030	.80	.1657	.1580	.90
416	.0812	.0885	.10	.0793	.0823	.50	.0818	.0962	.40	.1116	.1175	.40	.1427	.1189	.80
417	.1034	.0989	.40	.0842	.0805	.60	.1141	.0998	.50	.1617	.1320	.70	.1441	.1153	.90
All Items	.1040	.0937	.34	.1018	.0851	.39	.1014	.0980	.47	.1286	.1182	.51	.1803	.1299	.70

[1] Proportion of the random samples for which $T_1 > T_2$, that is for which the estimated beta distribution method was superior.

TABLE 4

ANOVA Summary Table for Index Values

Source	df	Mean Square	Mean Square Ratio	Estimated Variance Components[a]
Between	699			
Item	13	0.2112	76.32*	19.87
Sample Size (Size)	4	0.1884	36.73*	5.24
Item X Size	52	0.0051	1.85*	0.88
Sample/Item X Size	630	0.0028		
Within	700			
Method	1	0.0354	3.58	0.18
Item X Method	13	0.0099	15.67*	0.86
Size X Method	4	0.1258	10.06*	0.32
Item X Size X Method	52	0.0013	1.98*	0.23
Sample X Method/Item X Size	630	0.0006		
Total	1399			

* $p<.01$

[a]Actual values = entries X 10^{-4}

TABLE 5

Table of Means for the Method X Sample Size Interaction*

Method	Sample					All Samples
	.90(N≈170)	.75(N≈142)	.50(N≈94)	.25(N≈47)	.10(N≈19)	
Linear Interpolation	.1152	.1145	.1202	.1438	.1912	.1370
Beta	.1133	.1125	.1176	.1331	.1581	.1269
All Methods	.1142	.1135	.1189	.1384	.1747	

*Tukey critical difference for comparisons between any two means contained in the method X size interaction is .0112.

- 25 -

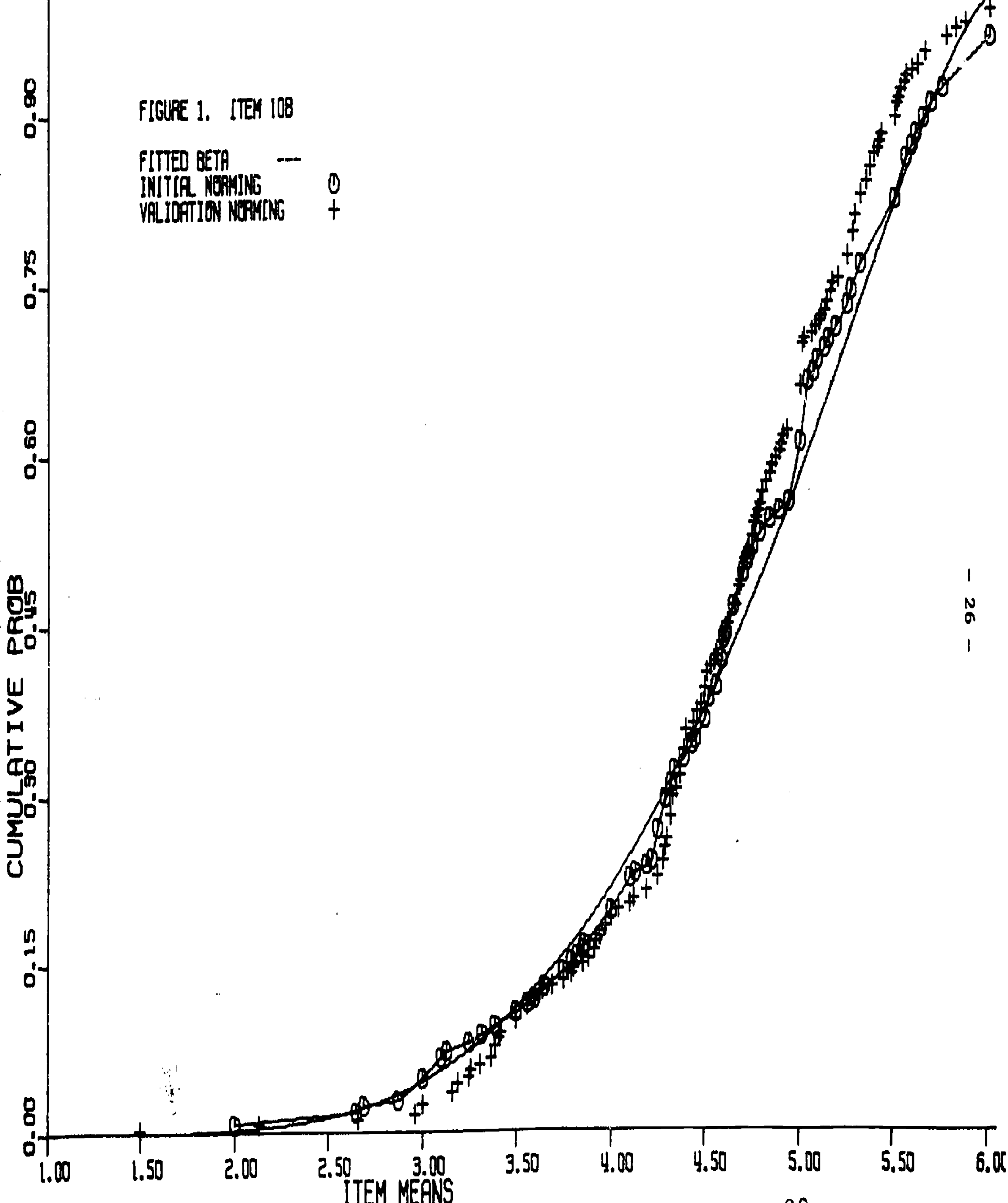

FIGURE 1. ITEM 10B
FITTED BETA
INITIAL NORMING
VALIDATION NORMING
CUMULATIVE PROB
0.00
0.15
0.30
0.45
0.60
0.75
0.90
1.00
1.50
2.00
2.50
3.00
3.50
4.00
4.50
5.00
5.50
6.00
ITEM MEANS

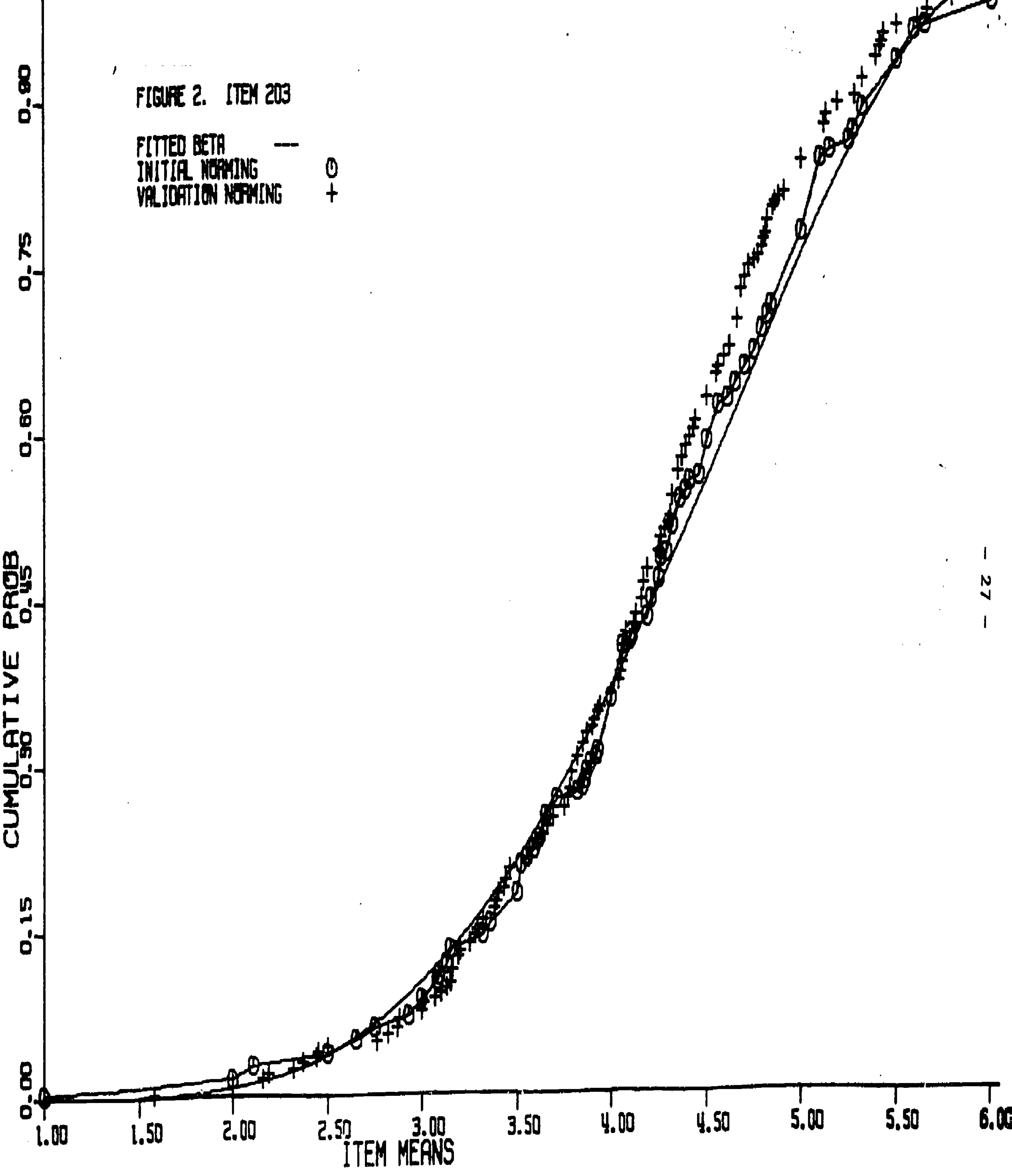
FIGURE 2. ITEM 203
FITTED BETA
INITIAL NORMING
VALIDATION NORMING
CUMULATIVE PROB
0.00
0.15
0.30
0.45
0.60
0.75
0.90
1.00
1.50
2.00
2.50
3.00
3.50
4.00
4.50
5.00
5.50
6.00
ITEM MEANS

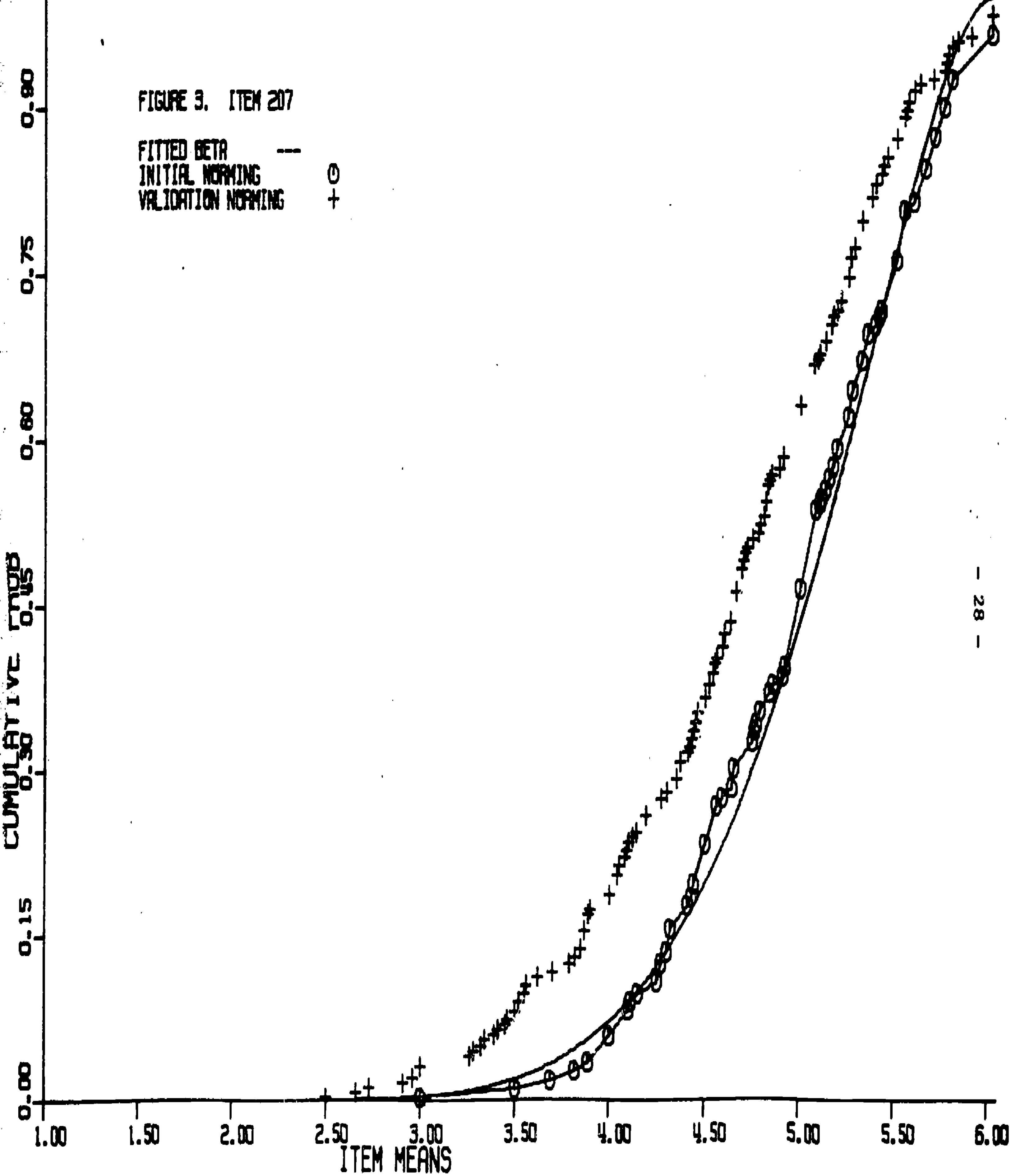
FIGURE 3. ITEM 207
FITTED BETA ---
INITIAL NORMING O
VALIDATION NORMING +
CUMULATIVE PROB
0.00
0.15
0.30
0.45
0.60
0.75
0.90
1.00
1.50
2.00
2.50
3.00
3.50
4.00
4.50
5.00
5.50
6.00
ITEM MEANS

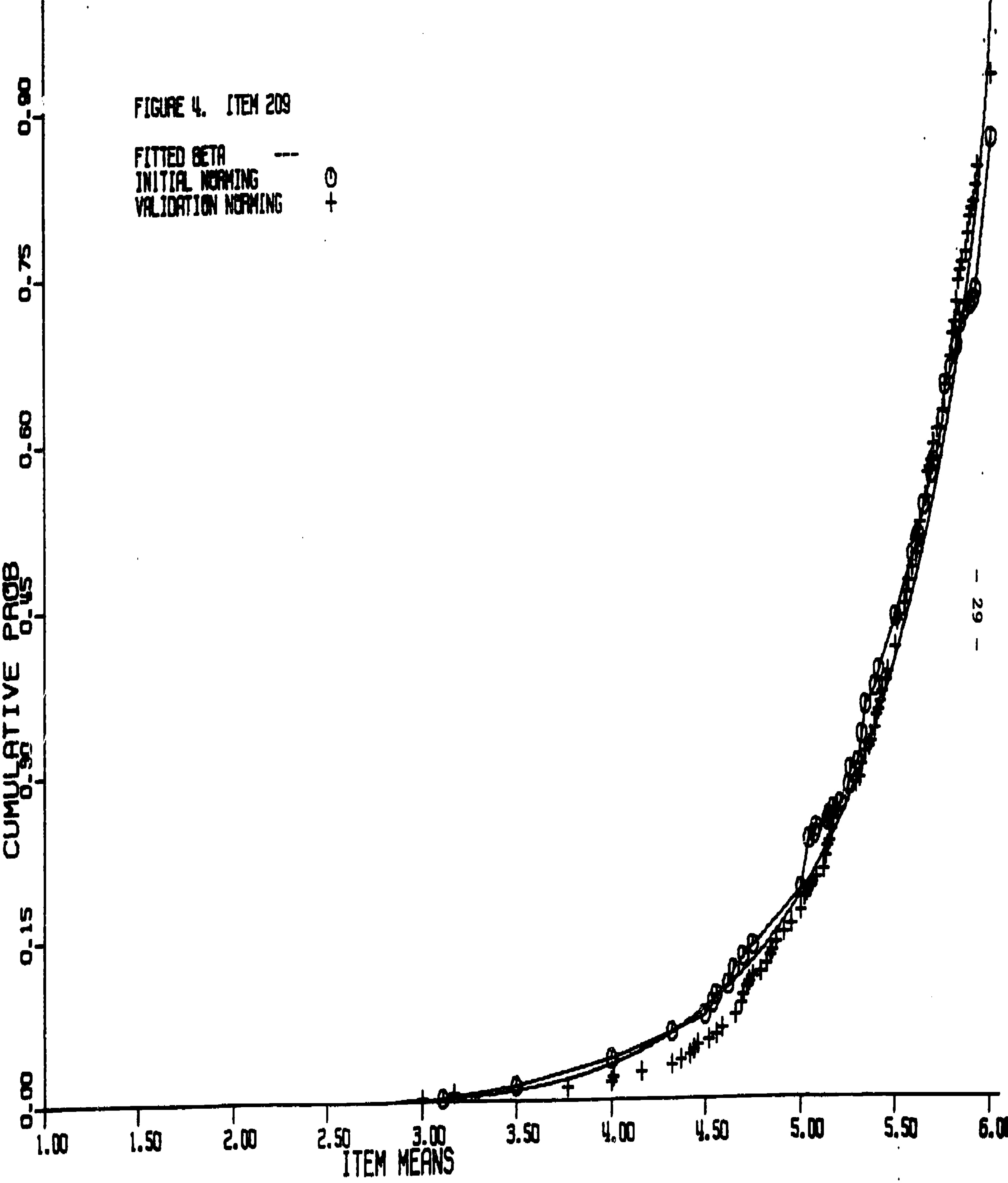
FIGURE 4. ITEM 209
FITTED BETA ---
INITIAL NORMING ⊖
VALIDATION NORMING +
CUMULATIVE PROB
0.00
0.15
0.30
0.45
0.60
0.75
0.90
1.00
1.50
2.00
2.50
3.00
3.50
4.00
4.50
5.00
5.50
6.00
ITEM MEANS

- 29 -

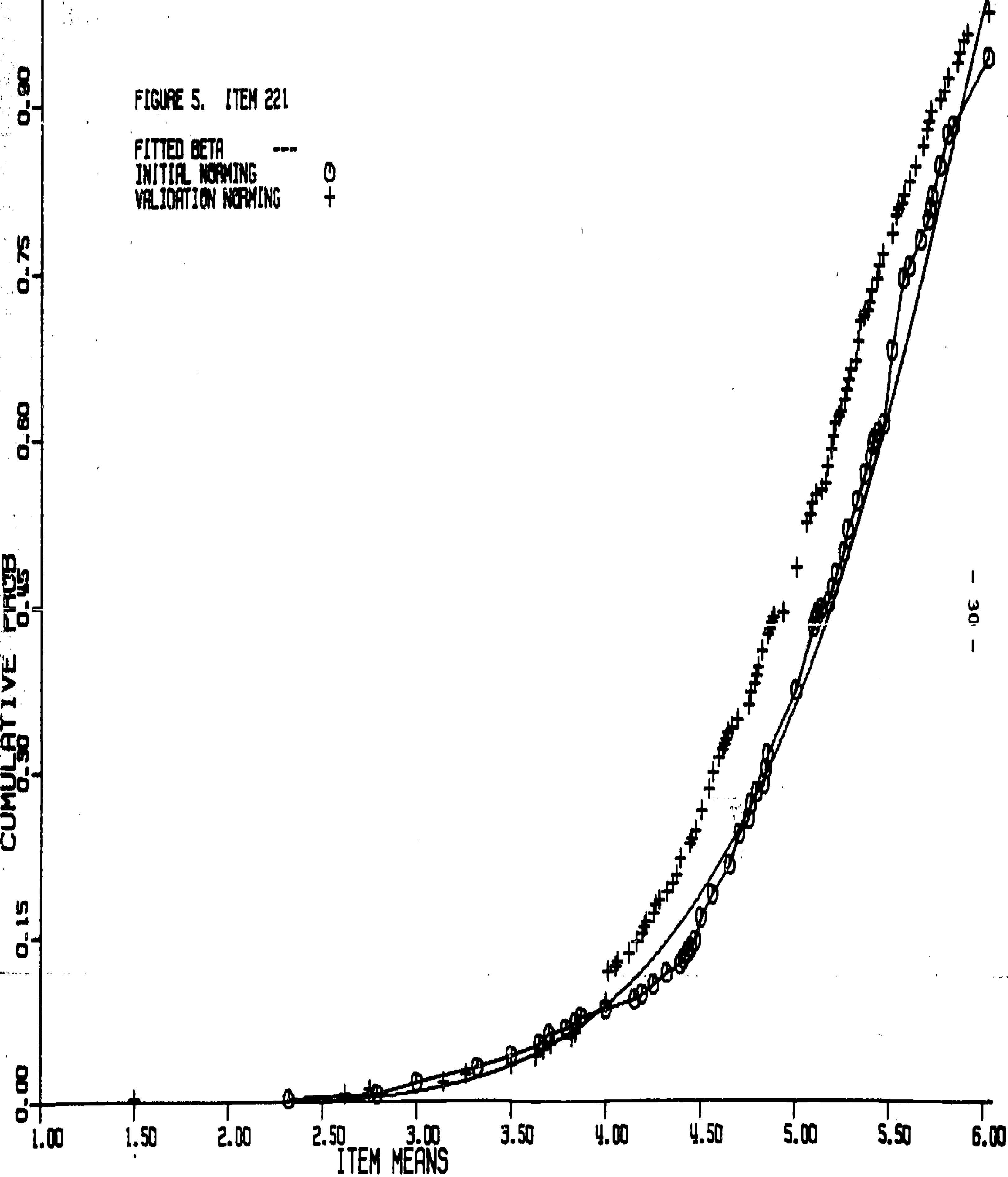

FIGURE 5. ITEM 221
FITTED BETA ---
INITIAL NORMING O
VALIDATION NORMING +
CUMULATIVE PROB
0.00
0.15
0.30
0.45
0.60
0.75
0.90
1.00
1.50
2.00
2.50
3.00
3.50
4.00
4.50
5.00
5.50
6.00
ITEM MEANS

- 30 -

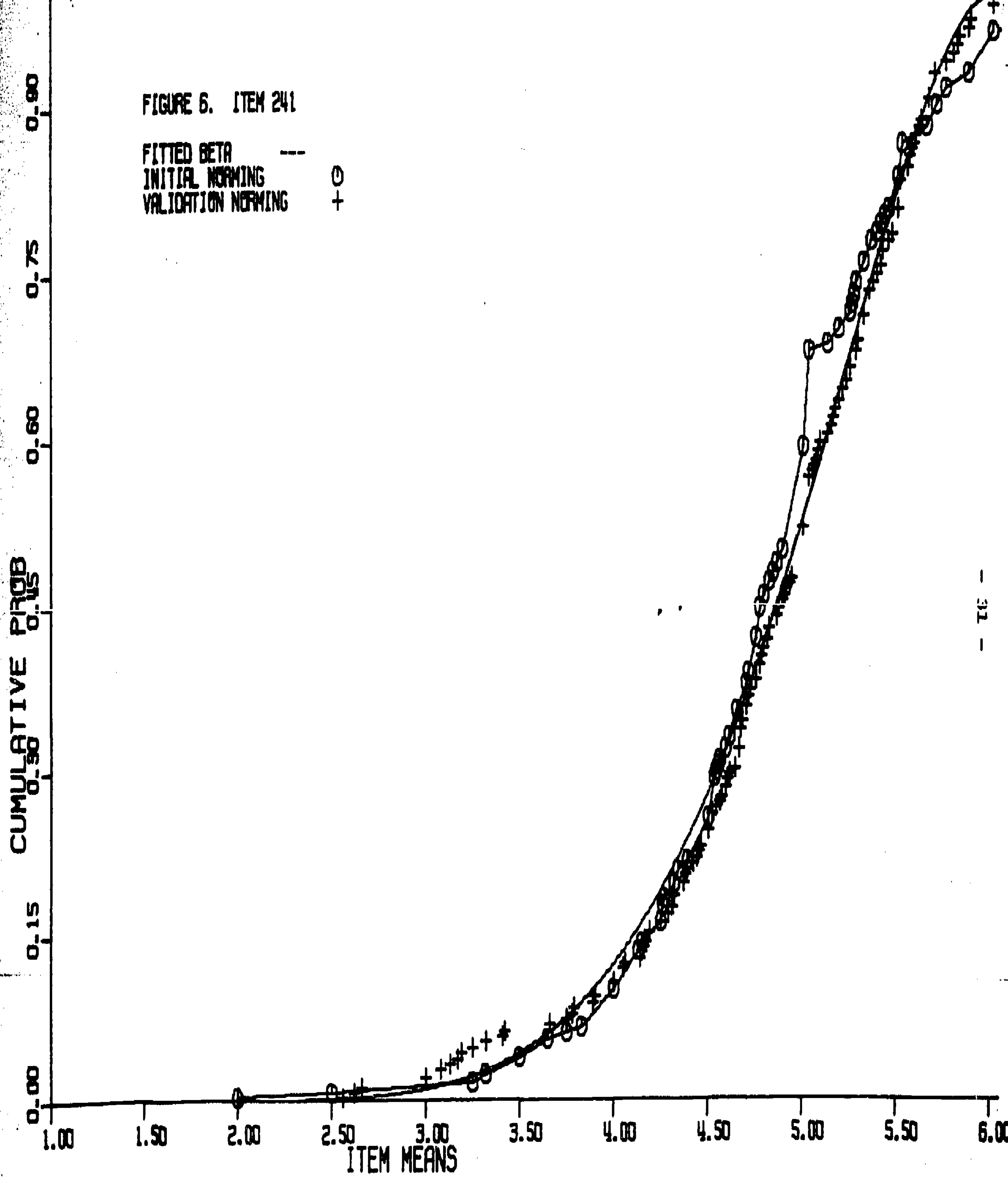
FIGURE 6. ITEM 241
FITTED BETA ---
INITIAL NORMING
VALIDATION NORMING +
CUMULATIVE PROB
0.00
0.15
0.30
0.45
0.60
0.75
0.90
1.00
1.50
2.00
2.50
3.00
3.50
4.00
4.50
5.00
5.50
6.00
ITEM MEANS

- 31 -

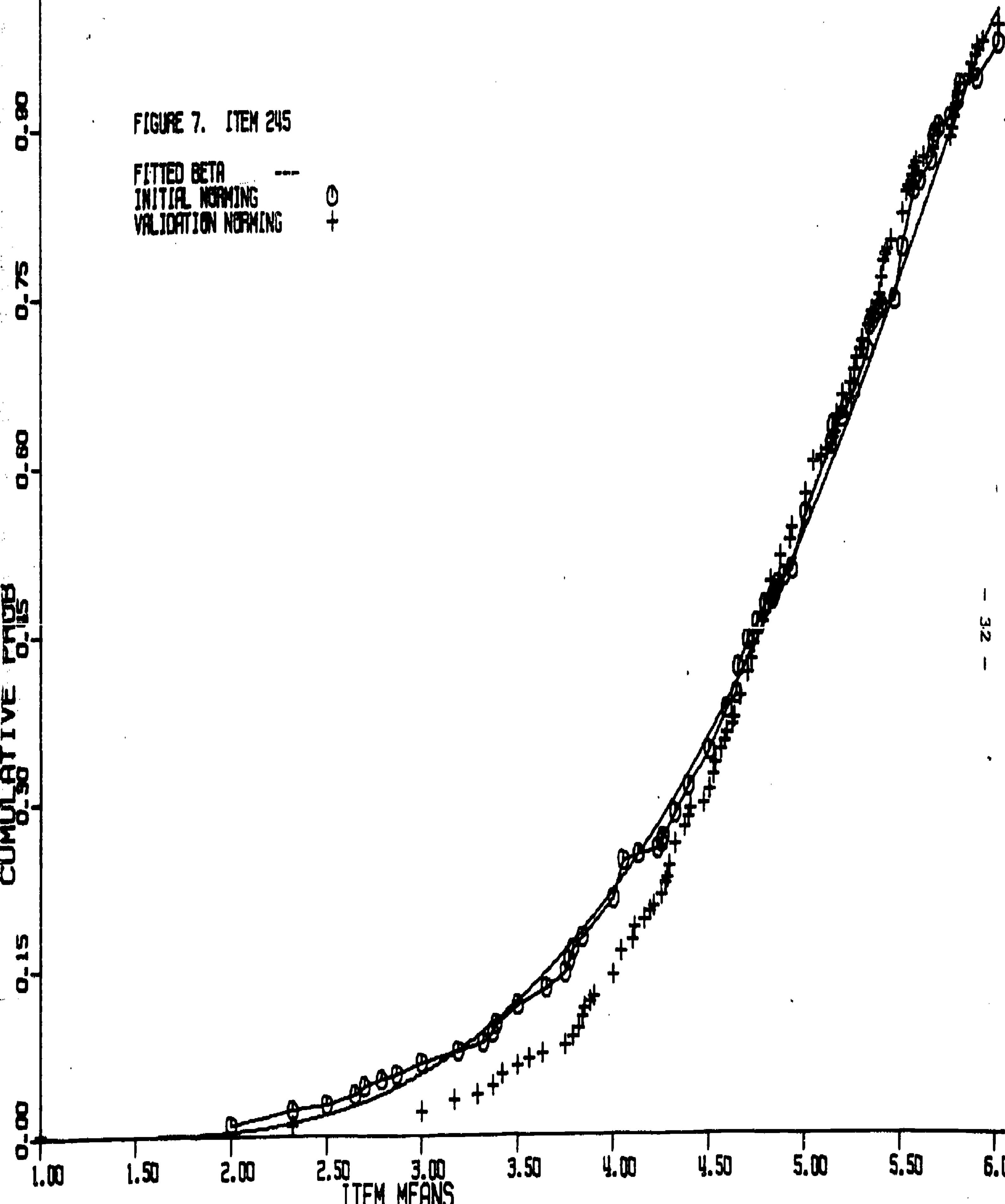

FIGURE 7. ITEM 245
FITTED BETA ---
INITIAL NORMING O
VALIDATION NORMING +
CUMULATIVE PROB
0.00
0.15
0.30
0.45
0.60
0.75
0.90
1.00
1.50
2.00
2.50
3.00
3.50
4.00
4.50
5.00
5.50
6.00
ITEM MEANS

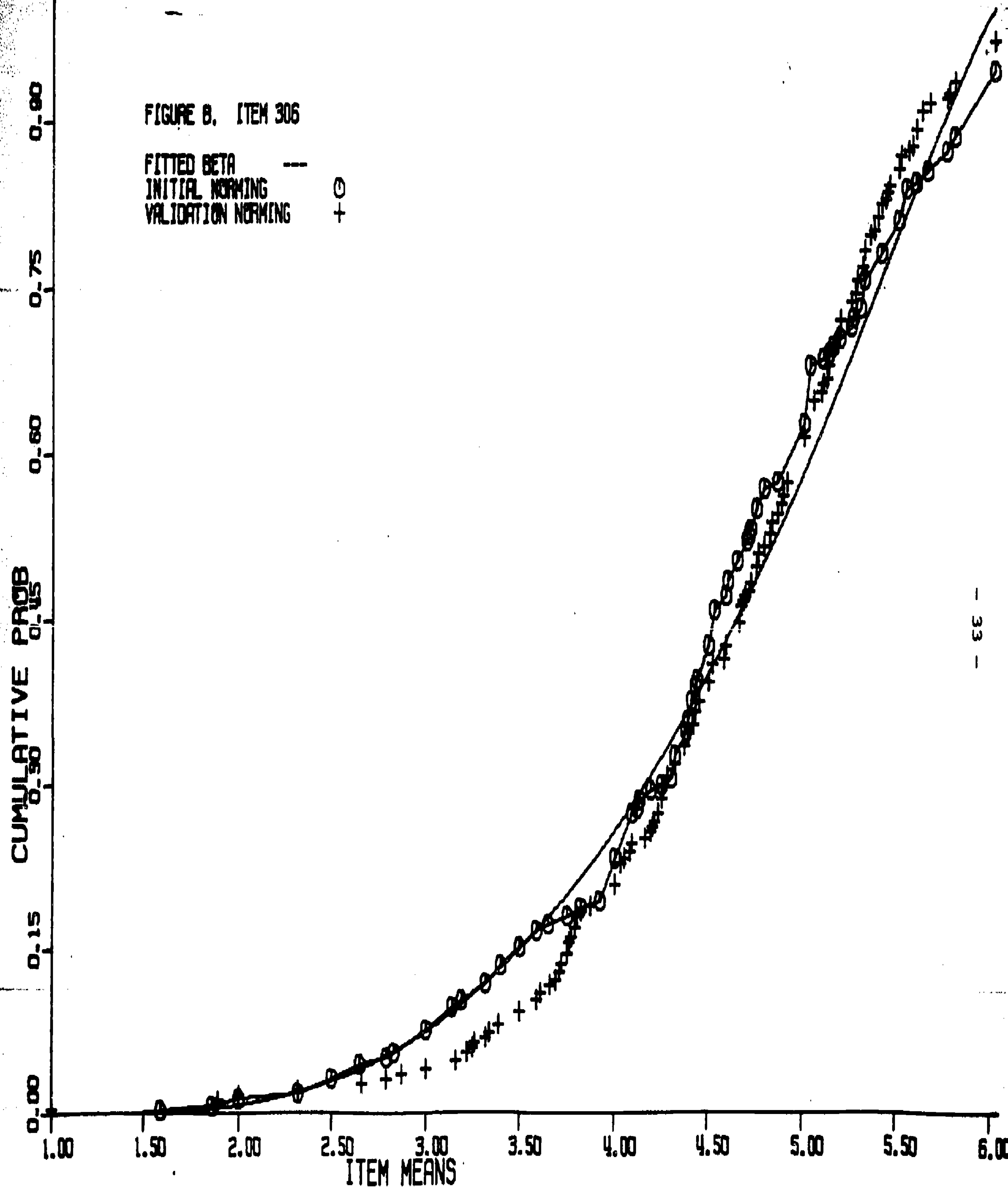
FIGURE 8. ITEM 306
FITTED BETA ---
INITIAL NORMING O
VALIDATION NORMING +
CUMULATIVE PROB
0.00
0.15
0.30
0.45
0.60
0.75
0.90
1.00
1.50
2.00
2.50
3.00
3.50
4.00
4.50
5.00
5.50
6.00
ITEM MEANS

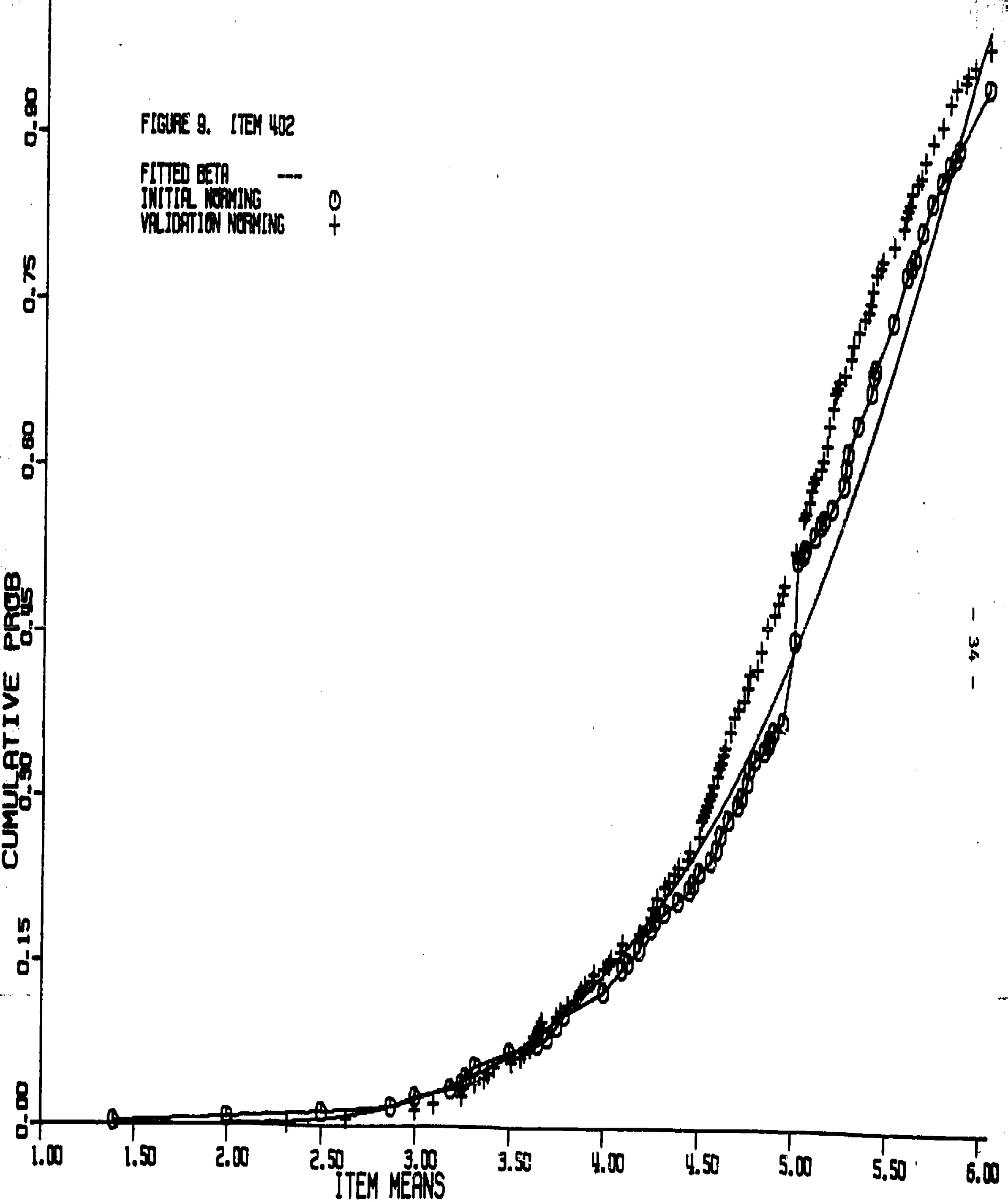
FIGURE 9. ITEM 402
FITTED BETA ----
INITIAL NORMING O
VALIDATION NORMING +
CUMULATIVE PROB
0.00
0.15
0.30
0.45
0.60
0.75
0.90
1.00
1.50
2.00
2.50
3.00
3.50
4.00
4.50
5.00
5.50
6.00
ITEM MEANS

- 34 -

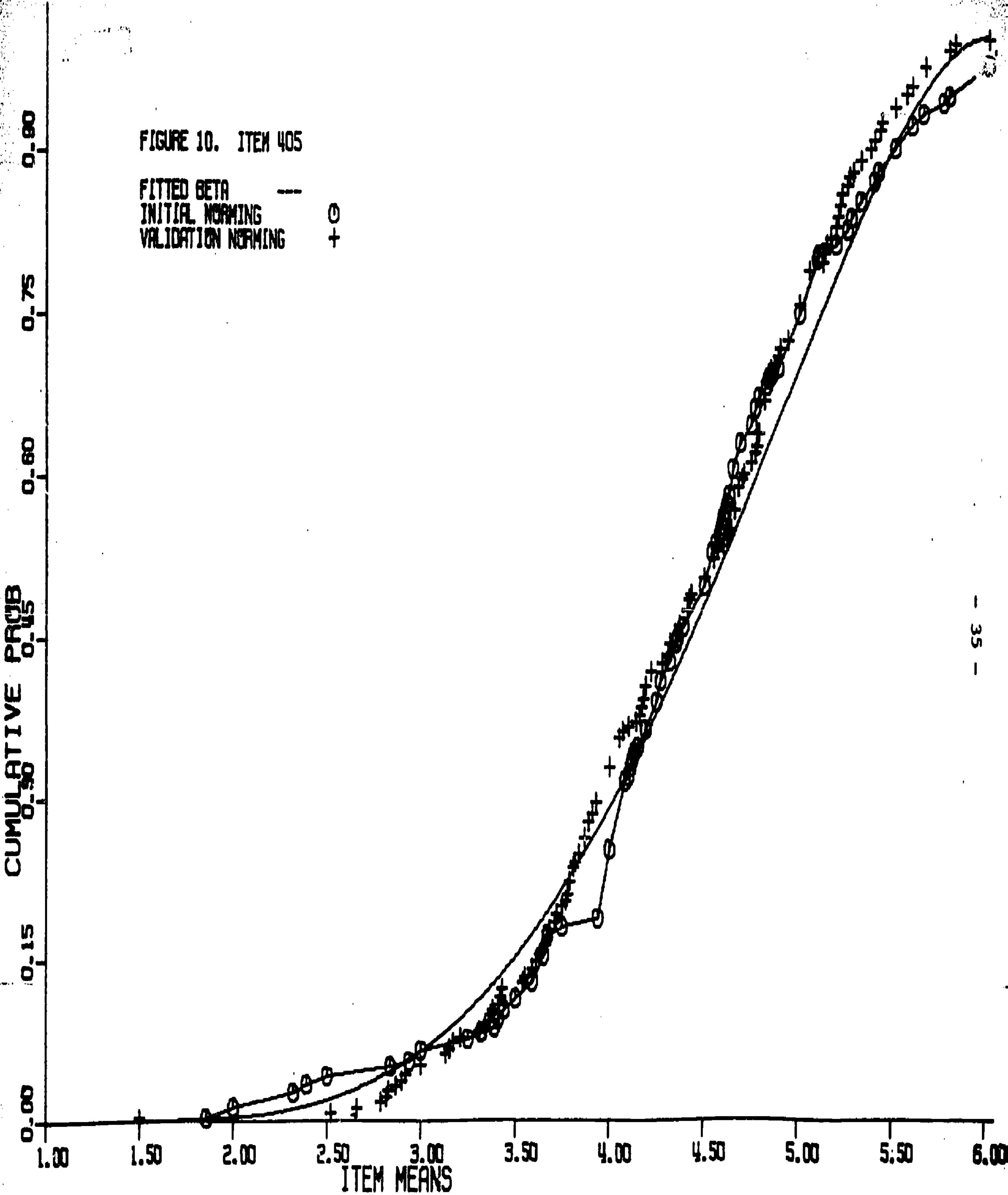
FIGURE 10. ITEM 405
FITTED BETA
INITIAL NORMING
VALIDATION NORMING
CUMULATIVE PROB
0.00
0.15
0.30
0.45
0.60
0.75
0.90
1.00
1.50
2.00
2.50
3.00
3.50
4.00
4.50
5.00
5.50
6.00
ITEM MEANS

- 35 -

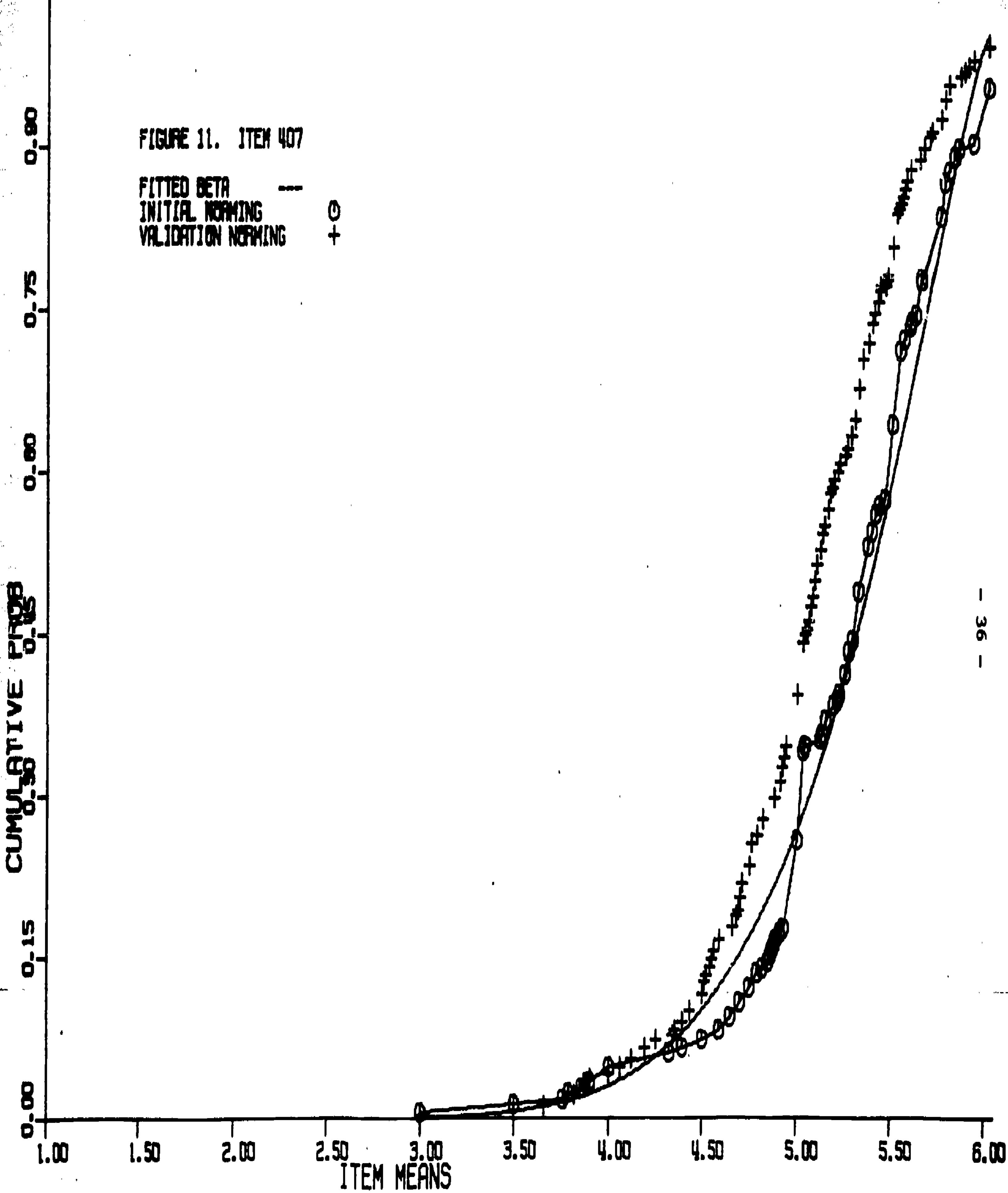

FIGURE 11. ITEM 407
FITTED BETA ---
INITIAL NORMING O
VALIDATION NORMING +
CUMULATIVE PROB
0.00
0.15
0.30
0.45
0.60
0.75
0.90
ITEM MEANS
1.00
1.50
2.00
2.50
3.00
3.50
4.00
4.50
5.00
5.50
6.00

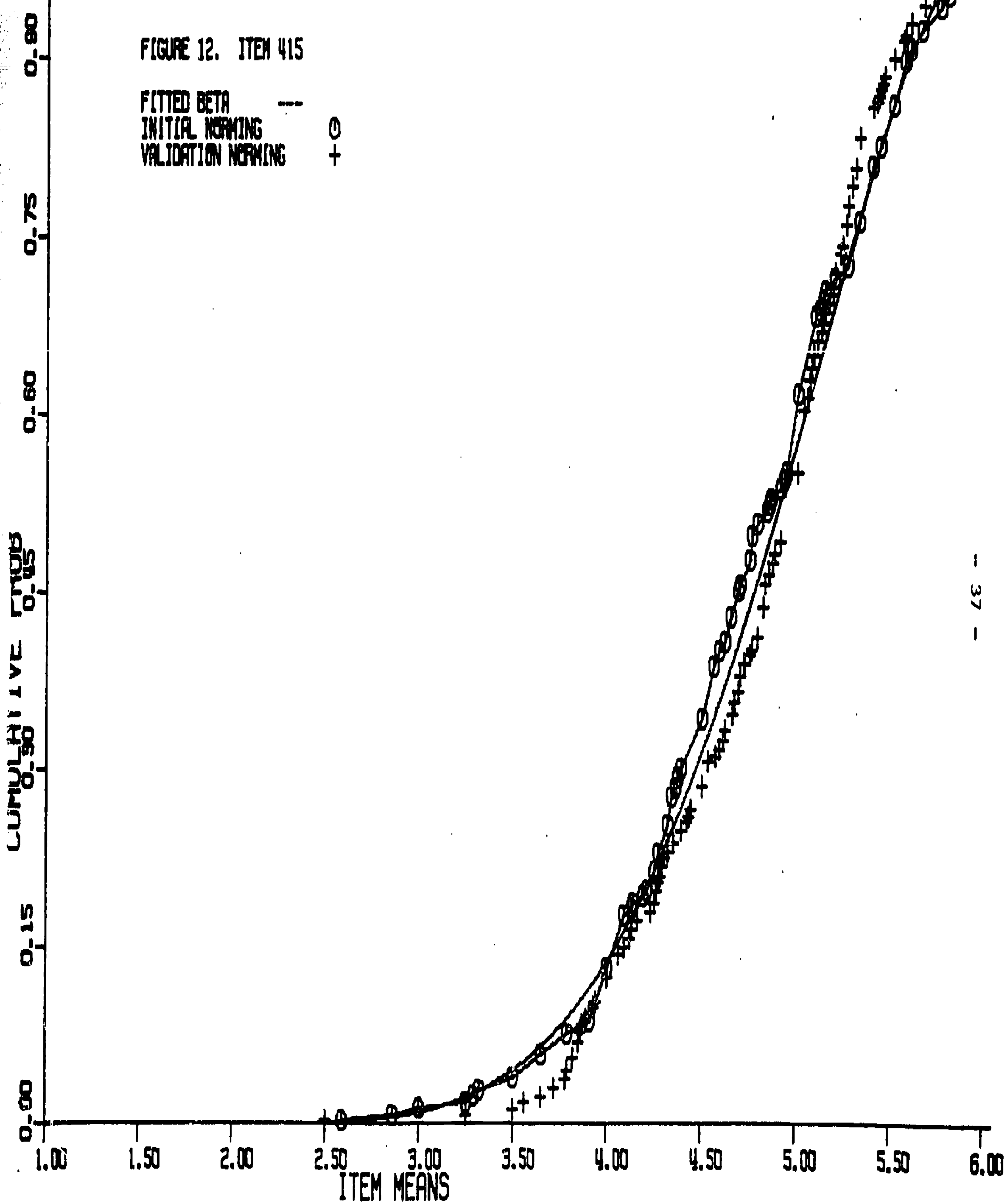

FIGURE 12. ITEM 415

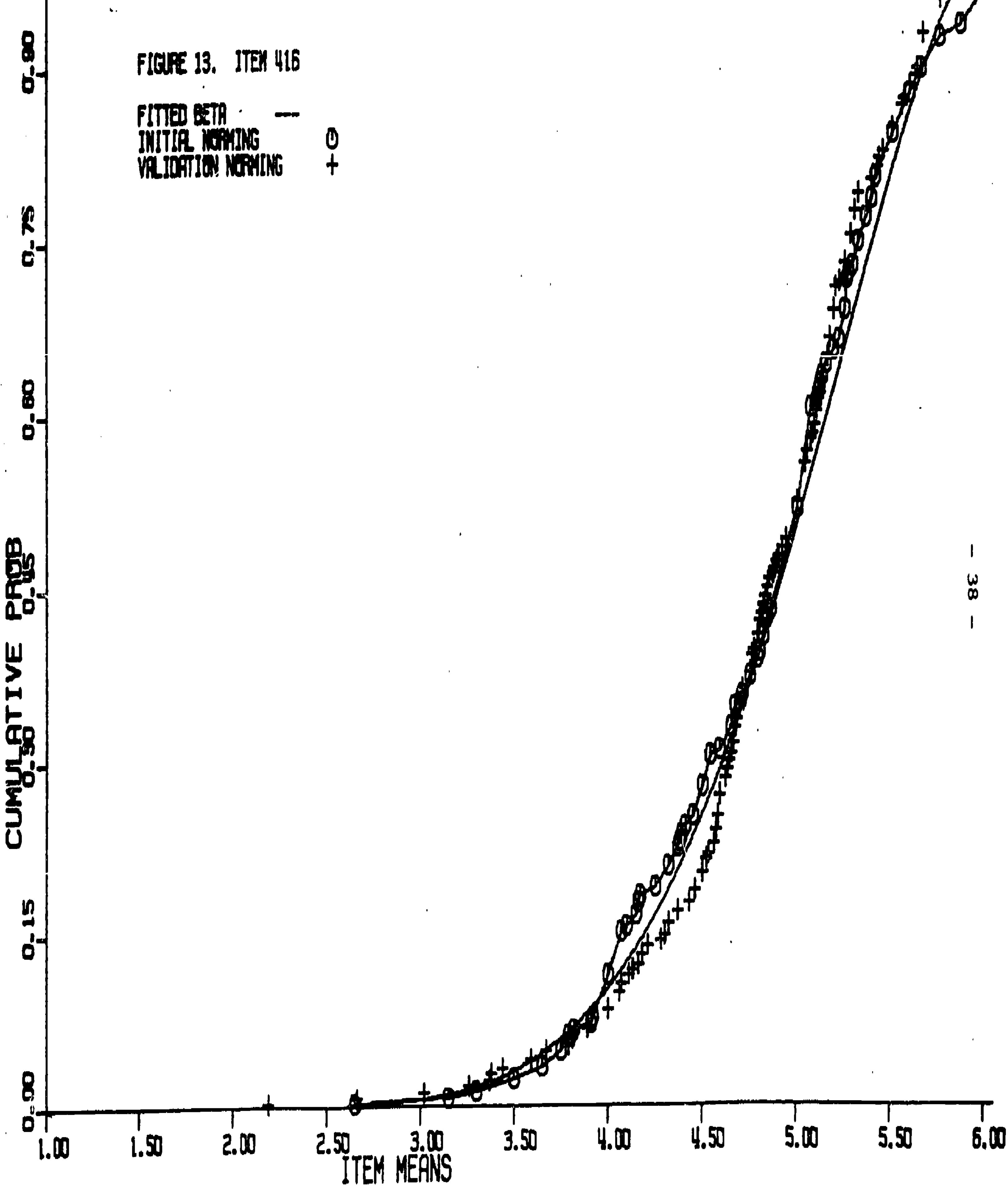
FIGURE 13. ITEM 416
FITTED BETA ---
INITIAL NORMING O
VALIDATION NORMING +
CUMULATIVE PROB
0.00
0.15
0.30
0.45
0.60
0.75
0.90
1.00
1.50
2.00
2.50
3.00
3.50
4.00
4.50
5.00
5.50
6.00
ITEM MEANS

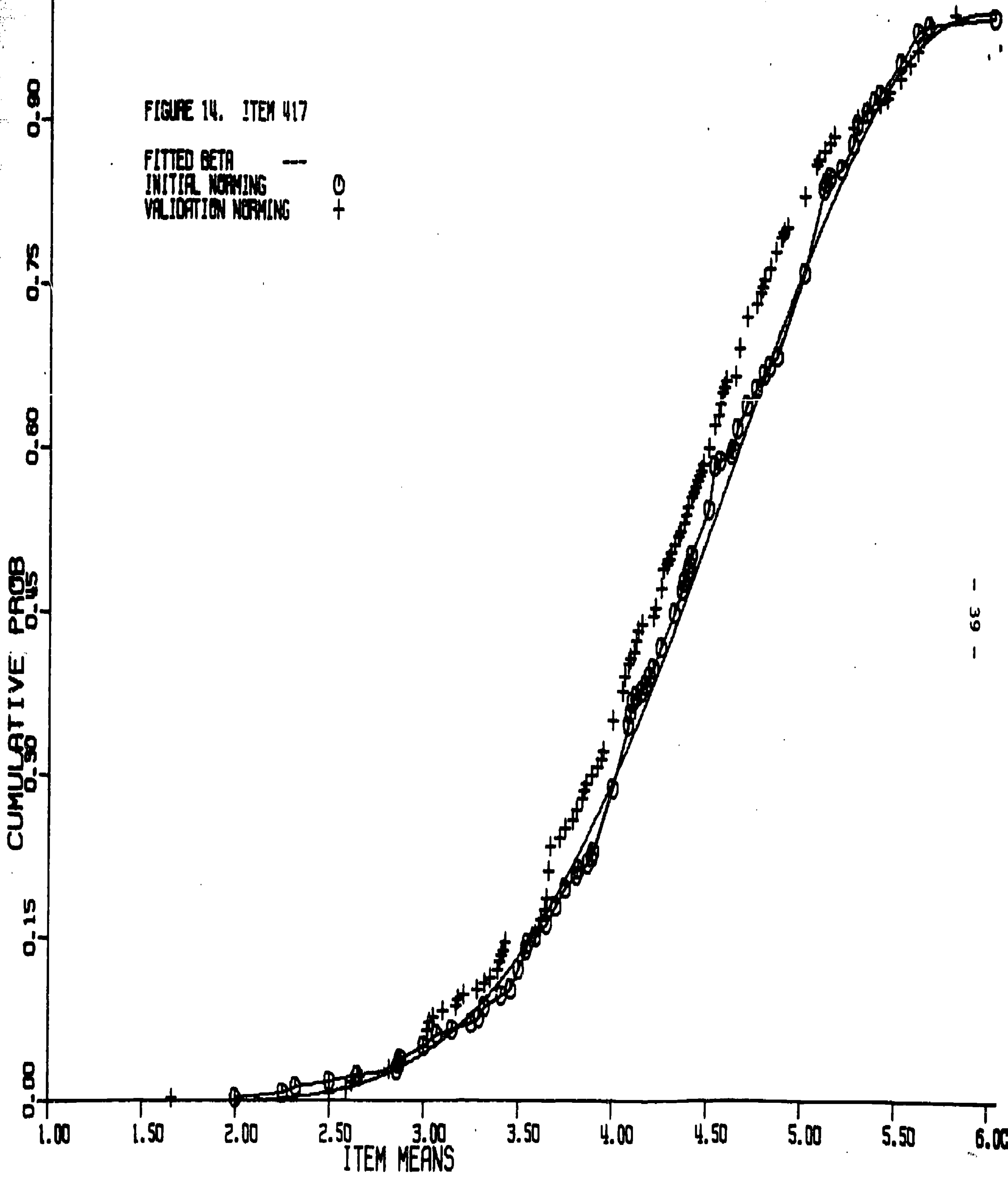
FIGURE 14. ITEM 417
FITTED BETA ---
INITIAL NORMING O
VALIDATION NORMING +
CUMULATIVE PROB
0.00
0.15
0.30
0.45
0.60
0.75
0.90
1.00
1.50
2.00
2.50
3.00
3.50
4.00
4.50
5.00
5.50
6.00
ITEM MEANS

- 39 -

Appendix

FORTRAN FUNCTION BCDF

Function BCDF(X) evaluates the integral of the beta distribution with parameters A & B to produce

$$BCDF(X) = \int_0^x B(A,B)y^{A-1}(1-y)^{B-1}dy.$$

The A and B parameters are passed to BCDF in the common statement

COMMON/BCDFC/RESET,A,B.

Since there is a fair amount of computation involved when A & B change values, this program will perform those computations only if RESET =.TRUE.. (Thus, RESET should equal .TRUE. for the first call.) After each evaluation of the beta, RESET will be automatically set to FALSE.. When A and/or B change values, the user should set RESET = .TRUE.. Note that this routine requires that the variable, y, be re-scaled to a zero-one scale. A listing of the program follows.

```
      FUNCTION BCDF(X)
C     FUNCTION BCDF WAS WRITTEN BY WILLIAM M. SMITH
C     APPROXIMATIONS FROM ABRAMOWICZ AND STEGUN(1965) HANDBOOK OF
C     MATHEMATICAL FUNCTIONS.
C     COMPUTES THE INCOMPLETE BETA INTEGRAL AT X WITH PARAMETERS A,B.
C
C     TO USE, REPLACE A AND B IN THE COMMON BLOCK WITH APPROPRIATE
C     PARAMETERS AND SET RESET TO .TRUE. PRIOR TO THE FIRST CALL.
C
      COMMON/BCDFC/RESET,AA,BB
C     THE FOLLOWING PARAMETERS ARE COMPUTED ON THE FIRST CALL. AND
C     ARE NOT NECESSARY TO COMPUTE THEREAFTER...
      LOGICAL RESET
      DATA IND,A1,A2,A3,C,D1,D2,D3,D4/0,8*0.0/,CUTOFF/0./
C     CHECK FOR FIRST CALL...
      IF(.NOT.RESET) GO TO 10
      A=AA
      B=BB
      IND=0
C     IF EITHER PARAMETER IS <2, USE RECURSION FORMULA...
      IF(A.LT.2.)GO TO 30
      IF(B.GT.2.)GO TO 20
```

```
30    IND=1
      D1=EXP(ALGAMA(A+B)-ALGAMA(A+1.)-ALGAMA(B))
      D2=(A+B)/(A+1.)
      D3=(A+B+1.)/B
      D4=(A+B+2.)/(B+1.)
      A=A+2.
      B=B+2.
20    CONTINUE
      AM1=A-1.
      BM1=B-1.
      ABM2=AM1+BM1
      CUTOFF=AM1/ABM2
      A1=.6666667*(B-A)/SQRT(ABM2*AM1*BM1)
      A2=8.333333E-2*(1./AM1+1./BM1-13./(A+BM1))
      A3=-5.333333E-1*A1*(A2+3./ABM2)
      C=AM1*ALOG(AM1)+BM1*ALOG(BM1)-ABM2*ALOG(ABM2)
      RESET=.FALSE.
10    IF(X.LE.0.)GO TO 60
      IF(X.GE.1.)GO TO 70
      Y=(2.*(C-AM1*ALOG(X)-BM1*ALOG(1.-X)))
      IF(Y.LT.0.)Y=0.
      Z=SQRT(Y)
      IF(X.LT.CUTOFF)Z=-Z
      BCDF=0.5*(ERF(.707107*Z)+1.)
      BCDF=BCDF-.3989423*EXP1(-.5*Y)*(A1+(A2*(Z-A1)+A3*(1.+.5*Y))/(1.+A2
     1 ))
      IF(IND.EQ.0)GO TO 55
      BCDF=BCDF+D1*X**(A-2.)*(1.-X)**(B-2.)*(1.+D2*X*(1.-D3*X*(1.+D4*(1.
     1 -X))))
55    IF(BCDF.LT.0.) BCDF=0.
      IF(BCDF.GT.1.) BCDF=1.
      RETURN
60    BCDF=0.
      RETURN
70    BCDF=1.
      RETURN
      END


      FUNCTION EXP1(X)
      EXP1=0.
      IF(X.GE.-70)EXP1=EXP(X)
      RETURN
      END
```